

NUCLEAR RIBOSOMAL DNA REVEALS INFRASPECIFIC VARIATION IN *CORETHROGYNE FILAGINIFOLIA* (ASTERACEAE)

JASPREET KAUR AND JYOTSNA SHARMA

Department of Plant and Soil Science, Texas Tech University, Lubbock, TX 79409
jaspreet.kaur@ttu.edu

ABSTRACT

Corethrogyne DC. (Asteraceae) is a controversial genus due to complex and confounding variability, and the subject of multiple taxonomic revisions. It was previously included within *Lessingia* Cham. The single species, *C. filaginifolia* Nutt., has high infraspecific variability. This study was conducted to better understand this infraspecific variability for *C. filaginifolia* var. *incana* (Lindl.) Canby and *C. filaginifolia* var. *linifolia* H. M. Hall. Barcoding loci from nuclear ribosomal DNA (nrDNA) and chloroplast DNA (cpDNA) were used along with morphological comparisons of voucher specimens. Based on two nrDNA loci, maximum parsimony, maximum likelihood, and Bayesian analyses, we detected five haplotypes of *C. filaginifolia*. The two focal varieties of this study, *C. filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia*, were indistinguishable at both, nrDNA and cpDNA loci.

Key Words: Asteraceae, California, DNA barcoding, *Lessingia*.

Asteraceae is one of the largest families of Angiosperms with more than 23,000 species (Panero and Funk 2008). Rapid diversification approximately 30 million years ago resulted in high morphological, chemical, and biological diversity (Jensen and Palmer 1987). Phylogenetic relationships at the tribal and subfamily levels within the Asteraceae have been examined extensively (Funk et al. 2009), though confusion still persists for some major lineages (Panero and Funk 2008). *Corethrogyne* DC., commonly known as sandaster, is one of several controversial genera in the Asteraceae. Its distribution is restricted to parts of the western United States (California and southwestern Oregon) and Mexico (northern Baja California) (Markos and Strother 2006; Baldwin et al. 2012) (Fig. 1a).

The taxonomy of *Corethrogyne* has been revised frequently. In 1959, Keck first proposed three species in the genus. By 1960, however, Ferris suggested seven species. Munz followed in 1974 to clarify the taxonomy of *C. filaginifolia* Nutt. using only plant material from southern California, and presented 11 varieties. Approximately two decades later, Lane (1992, 1993) unified the genus *Corethrogyne* with *Lessingia* Cham. based on morphological and chloroplast DNA similarities. Between 1993 and late 1990s, various studies (Zhang and Bremer 1993; Nesom 1994; Morgan 1997) showed that *Lessingia* is closely related to *Corethrogyne* based on morphology and chloroplast DNA. More recently, however, Saroyan et al. (2000) published a revision presenting *Corethrogyne* as distinct from *Lessingia* based on the perennial life cycle and radiate inflorescence heads in *Corethrogyne* as opposed to an annual life cycle and sometimes discoid inflorescence in *Lessingia*. Notably, the same authors also reported morphometric distinctiveness between northern and southern populations of *C. filaginifolia* in California based on the

number of inflorescences per flowering stem and length:diameter ratio of the involucre. Subsequently, in 2001, Markos and Baldwin provided genetic evidence to support the discrimination between *Lessingia* and *Corethrogyne*. In 2012, Markos and Strother recognized a single species of *Corethrogyne*, *C. filaginifolia* Nutt., without any infraspecific taxa. Simultaneously, California Native Plant Society (CNPS 2015) maintains two species within the genus *Corethrogyne* (*C. leucophylla* Jeps. and *C. filaginifolia*) and recognizes two varieties of *C. filaginifolia*: 1) *C. filaginifolia* Nutt. var. *incana* (Lindl.) Canby, and 2) *C. filaginifolia* Nutt. var. *linifolia* H.M. Hall. Adding to the confusion further, CalFlora (2015) also recognizes the same two species (*C. leucophylla* and *C. filaginifolia*), but maintains four varieties within *C. filaginifolia*: 1) *C. filaginifolia* Nutt. var. *californica* (DC.) Saroyan, 2) *C. filaginifolia* Nutt. var. *filaginifolia*, 3) *C. filaginifolia* var. *incana*, and 4) *C. filaginifolia* var. *linifolia*. The three *Corethrogyne* taxa recognized by CNPS are also listed in the rare, and endangered plant inventory of CNPS (2015), but considering the taxonomic uncertainty within the genus and within the *C. filaginifolia* complex in particular, the conservation status of these taxa remains unclear (Markos and Baldwin 2001).

Two of the recognized taxa, *C. filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia*, are given a rarity rank of 1.B1 (CNPS 2015). Both taxa are endemic to San Diego County and occur in coastal scrub and chaparral, but occupy different areas within this narrow geographic range (Munz 1974; CNPS 2015) (Fig. 1b). *Corethrogyne filaginifolia* var. *incana* is restricted to sandy soils of the Point Loma peninsula. *Corethrogyne filaginifolia* var. *linifolia* occurs on open bluffs and in shrubby habitats of Del Mar.

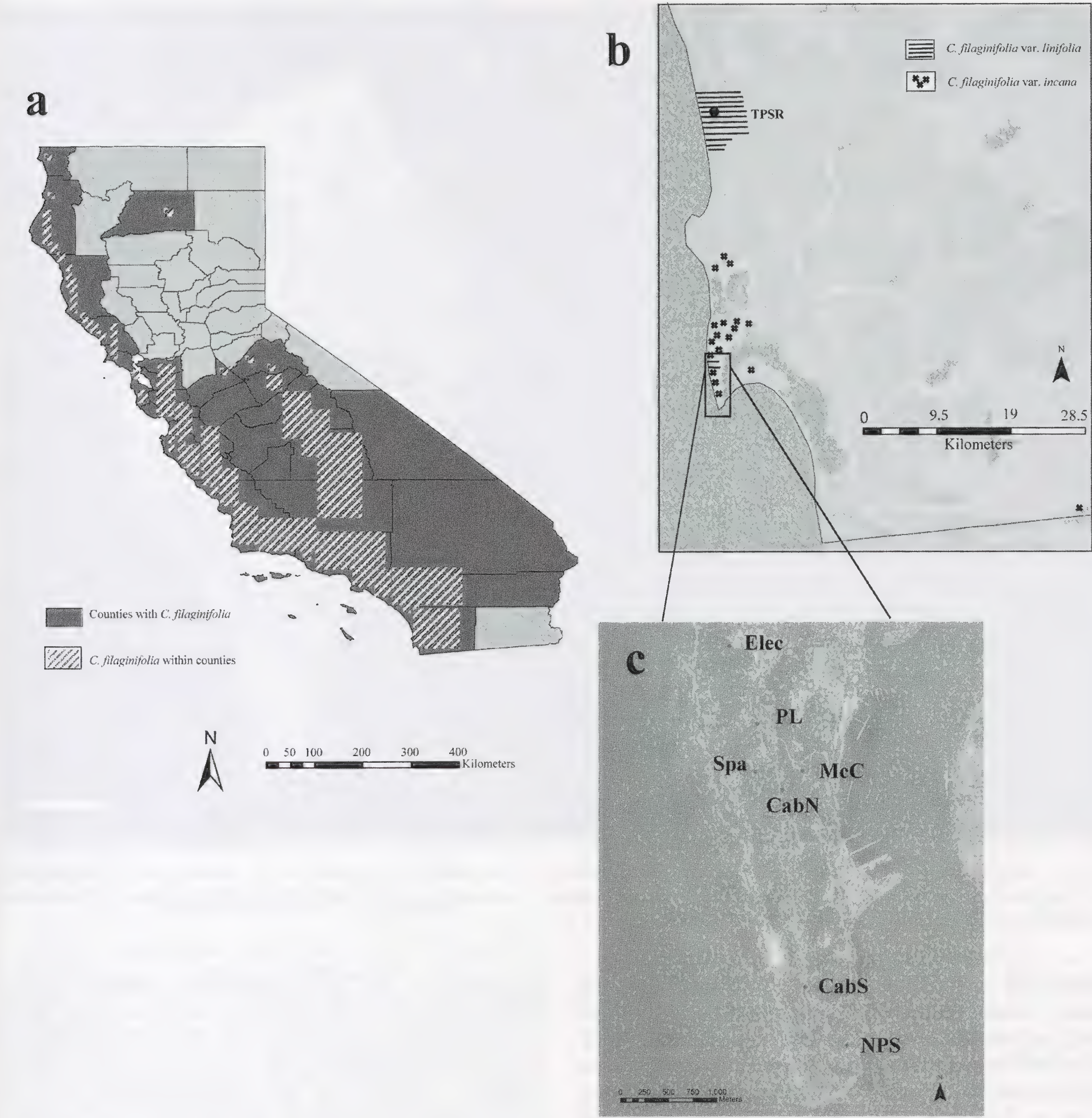


FIG. 1. Maps showing the distribution of *Corethrogyne filaginifolia* and locations where individuals were sampled for this study. a) A map of California showing the counties hosting *C. filaginifolia* populations, and coarse distribution of *C. filaginifolia* within them indicated by hatching (CalFlora 2015). b) A map of San Diego County showing the approximate distribution of *C. filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia* (also from CalFlora 2015). Additionally, a sampling location within Torrey Pine State Reserve (TPSR) is shown where *C. filaginifolia* var. *linifolia* was sampled. c) A map of Point Loma peninsula showing locations where the two varieties of *C. filaginifolia* were sampled; Cabrillo Drive (CabN, CabS), McClelland Road (McC), SPAWAR (Spa), Electron Drive (Elec), and National Park Service (NPS) represent the sampling sites for *C. filaginifolia* var. *incana*, while the location coded as Point Loma (PL) represent the sampling site for *C. filaginifolia* var. *linifolia*.

However, both taxa have been reported (as collections and occurrences) from outside these core ranges (CalFlora 2015). CNPS (2015) reported that the distribution of *C. filaginifolia* var. *linifolia* extended to Rancho Santa Fe, La Jolla, Del Mar, Encinitas and San Luis Rey, and the distribution of *C. filaginifolia* var. *incana* included National City, Point Loma, Imperial Beach, Del Mar, La Jolla, and Rancho Santa Fe (Fig. 1b). Because both taxa (*C. filaginifolia* var.

incana and *C. filaginifolia* var. *linifolia*) are present on the Point Loma peninsula where they may potentially hybridize, conservation concerns arose with respect to their taxonomic uniqueness.

Corethrogyne filaginifolia var. *incana* can be distinguished morphologically from *C. filaginifolia* var. *linifolia* by presence of distinct stout-stalked glands (Fig. 2a) on the involucral bracts and distal plant parts making them sticky to touch. In comparison, *C.*

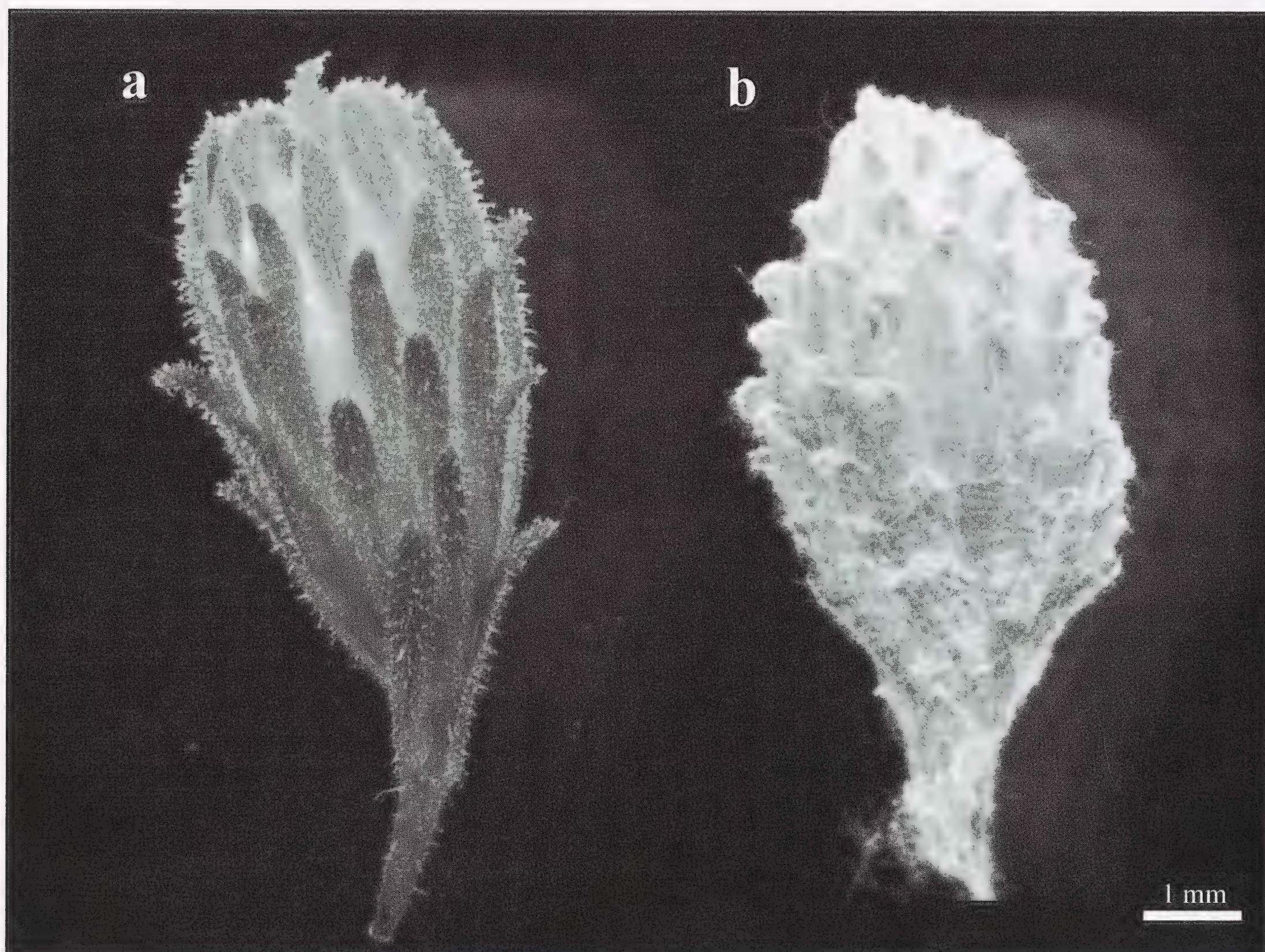


FIG. 2. Morphological differences between involucral bracts of *Corethrogyne filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia* are shown. a) Involucral bracts of *C. filaginifolia* var. *incana* are covered with stout-stalked glands. b) Involucral bracts of *C. filaginifolia* var. *linifolia* have tomentose surfaces.

filaginifolia var. *linifolia* has tomentose involucral bracts, lacking glands (Munz 1974) (Fig. 2b). Munz (1974) reported additional morphological and phenological differences to discriminate between the two taxa which included; 1) plant height (*C. filaginifolia* var. *incana* = 50–80 cm, *C. filaginifolia* var. *linifolia* = 20–40 cm), 2) involucre length (*C. filaginifolia* var. *incana* = 10–12 mm, *C. filaginifolia* var. *linifolia* = 8–10 mm), and 3) leaf shape (*C. filaginifolia* var. *incana* = linear to narrowly lanceolate, *C. filaginifolia* var. *linifolia* = narrow with 1–2 mm width). Further, genetic differences between the two varieties have not been investigated to test the taxonomic distinctiveness.

We sought to assess whether the two varieties, *Corethrogyne filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia*, could be distinguished using multiple DNA barcoding loci (Hebert et al. 2003). Toward this end, we compared the two taxa using six genomic loci that have previously been used to discriminate taxa within the Asteraceae and in other plant families (Chase et al. 2007; Kress et al. 2007; CBOL Plant Working Group 2009; Chen et al. 2010; Gao et al. 2010).

Further, to broadly compare with the Californian *C. filaginifolia* complex, we combined our data

generated in this study (without vouchered specimens) from the two varieties with sequence data (with vouchered specimens) previously generated by Markos and Baldwin (2001, 2002) and from NCBI (PopSet 417346869). We hypothesized that: 1) *Corethrogyne filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia* are genetically distinguishable from each other at the selected barcode loci. We based this hypothesis on the observations of Munz (1974) and our personal observation of the morphological differences between the two varieties; and 2) individuals within the Californian *C. filaginifolia* complex show sub-specific variation based on a latitudinal gradient. This hypothesis was based on the morphometric study of Saroyan et al. (2000), which suggested the distinction between northern and southern populations of *Corethrogyne* in California.

MATERIALS AND METHODS

Sample collection and DNA extraction

Three individuals of *C. filaginifolia* var. *incana* were sampled from each of the following six

locations: Cabrillo Drive (CabN, CabS), McClelland Road (McC), SPAWAR (Spa), Electron Drive (Elec), and National Park Service (NPS) on Point Loma peninsula, San Diego County (Fig. 1c). The sampled sites for *C. filaginifolia* var. *incana* were within approximately 0.5–5 km from each other.

Three individuals belonging to *Corethrogyne filaginifolia* var. *linifolia* were sampled at each of the following two locations: Point Loma (PL) (a central site on the peninsula) and Torrey Pines State Natural Reserve (TPSR) in Del Mar, San Diego County (Fig. 1b and c). TPSR is approximately 25–26 km from PL.

Up to 10 leaves were collected from each plant in May 2014 in order to obtain DNA for the barcode analyses. Plants were collected at reproductive stage when *C. filaginifolia* var. *incana* can be easily distinguished from *C. filaginifolia* var. *linifolia* by the presence of stout-stalked glands on its involucre bracts (Fig. 2). No plant materials were collected for herbarium voucher specimens. Leaf samples were stored at 4°C immediately upon collection.

Leaves from each individual were washed and sliced into 1–2 mm sections upon arrival at the laboratory in Texas Tech University, Lubbock, TX, to prepare them for DNA extraction. DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) was used to obtain DNA following the manufacturer's protocol. DNA quantity and quality was checked using NanoDrop ND1000 (Thermo Scientific, Waltham, MA, USA) and confirmed by running 1 µl genomic DNA on a 1% agarose gel.

PCR Amplification and Sequencing

For *C. filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia*, we used four cpDNA (*psbA-trnH*, *rbcL*, *rpoC1*, *ycf5*), and two nrDNA (ITS and ITS2) plant barcoding loci that have previously been used to discriminate taxa within the Asteraceae and in other plant families (Chase et al. 2007; Kress et al. 2007; CBOL Plant Working Group 2009; Chen et al. 2010; Gao et al. 2010). These loci were amplified via polymerase chain reaction (PCR) in a Mastercycler EPGradient (Eppendorf, Hamburg, Germany) using genomic DNA. Primer sequences and PCR thermal profiles were obtained from Chen et al. (2010). For each reaction, a 25 µl mixture was prepared by using the Promega GoTaq Flexi DNA Polymerase reagent kit (Promega, Madison, WI, USA). Reagent concentrations were 5 µl of 5× Green GoTaq Flexi Buffer, 0.4 µl of 40 mM dNTPs (10 mM of each dNTP), 4 µl of 25 mM MgCl₂, 1 µl of each primer (10 µM), 0.25 µl of 10 µg/µl bovine serum albumin, 0.25 µl of 5U/µl GoTaq DNA polymerase, 4 µl of 50–100 ng/µl DNA, 9.1 µl of sterile H₂O. During optimization, PCR was first performed by using the thermal profiles recommended by Chen et al. (2010). The loci that did not amplify or amplified inconsistently were exposed to touchdown PCR by adding a ±5°C range to the annealing temperature recommended by Chen et al.

(2010) for each primer. Amplicons were purified using DNA Clean and Concentrator 5 kit (Zymo Research, Irvine, CA, USA) and were sent for bi-directional sequencing in an ABI 3730xl genetic analyzer (Applied Biosystems, Foster City, CA, USA) at the DNA Analysis Facility on Science Hill, Yale University, New Haven, CT.

We then generated consensus sequences by using SeqTrace version 0.9.0 (Stucky 2012). Only reads with >100 bp length, and showing >50% overlap between forward and reverse primer reads were selected for further analyses. Quality trimming and base calling of the consensus sequences was subsequently performed using the same software. The ends of consensus sequences were trimmed if they had at least two nucleotide bases exhibiting < 20 Phred quality score (Q) using 20 bp window.

Reference Sequence Extraction from GenBank

Previously generated nrDNA sequences of 3'ETS (30 sequences), 5'ETS (5 sequences), and ITS loci (31 sequences) from vouchered plants of *C. filaginifolia* (Benson et al. 1999; Markos and Baldwin 2001, 2002; NCBI PopSet 417346869) were included in the phylogenetic analyses (Table 1). We also included 3'ETS and ITS sequences from *Lessingia arachnoidea* Greene and *L. glandulifera* A. Gray to represent genera closely related to *Corethrogyne* (Markos and Baldwin 2001; Table 1). Additionally, previously generated cpDNA sequences from *Bidens pilosa* were included to represent an outgroup for *psbA-trnH*, *rbcL*, *rpoC1*, and *ycf5* (Chen et al. 2010).

Sequence Alignment and Sequence Divergence

All sequences were aligned, and consistency scores for the alignments were estimated in T-Coffee version 11.00 (Notredame et al. 2000) with M-Coffee mode. The alignments of nrDNA and cpDNA loci were concatenated separately and partitioned into data blocks where each data block represented an individual locus. The protein coding loci *rbcL*, *rpoC1*, and *ycf5* were further partitioned based on the codon position. The partitioned datasets were used to predict the best model of DNA evolution in Partition Finder version 1.1.1 using Bayesian Information Criterion (BIC) score (Lanfear et al. 2012). Sequence divergences were determined for each aligned locus using Kimura 2-parameter (K2P) DNA substitution model with gamma distribution after pairwise deletion of gaps in MEGA (Molecular Evolutionary Genetics Analysis) version 6.0 (Tamura et al. 2013).

Phylogenetic Analyses

Maximum parsimony (MP) trees were generated in PAUP* test version 4.0a146 (Swofford 2003) with a heuristic search and tree bisection-reconnection (TBR) algorithm of branch swapping. Sequence addition was random, and all characters were given

an equal weight. The 'multrees' option was kept ineffective, and 1000 bootstrap replicates were used to calculate the clade support values. Subsequently, a 50% majority rule consensus tree was generated from the bootstrap replicates. Consistency index (CI) and Retention index (RI) were also calculated in PAUP*. Additionally, Maximum Likelihood (ML) analysis was conducted with RAxML-HPC2 version 7.4.2 (Stamatakis 2006) by generating 1000 ML trees using the random parsimonious tree as the starting tree and using GTRGAMMAI as the DNA substitution model. One thousand bootstrap replicates were used to estimate clade support on the maximum likelihood tree. Finally, we performed Bayesian analysis using MrBayes version 3.2.5 (Ronquist and Huelsenbeck 2003). The best DNA substitution models were selected based on Partition Finder results, and prior probabilities were set to the default values for selected models. Two runs were initiated with random trees, and four chains were included in each run of which three were heated and one was cold. Analysis was carried out for >1 million generations with the temperature set to 0.2 for sufficient swapping between hot and cold chains. Cold chains were sampled after every 10,000 generations and cross diagnosis of two runs was carried out after every 500 generations. Tracer version 1.6 (Rambaut and Drummond 2007) was used to determine the number of generations with consistent marginal probabilities. To compute the consensus Bayesian tree, results from initial 25% generations were discarded owing to their unstable posterior probability values. Fig Tree version 1.4.2 (Rambaut 2007) was used to visualize the phylogenetic trees rooted with their respective outgroups.

Here we report the results of MP, ML, and Bayesian analyses based on combined 1) GenBank extracted nrDNA loci sequences of *C. filaginifolia*, 2) cpDNA loci from *C. filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia* individuals.

Statistical Analyses

We used sequences downloaded from GenBank representing *C. filaginifolia* specimens from across California to conduct an analysis of molecular variance (AMOVA) to determine haplotype diversity in GENALEX version 6.5 (Peakall and Smouse 2006). The sequences were divided into four groups based on ecoregions and sub-ecoregions of California (McNab et al. 2007). Sequences representing Del Norte, Mendocino, and Marin counties were categorized as a single group from Northern California Coast region. Those from Santa Clara, Santa Cruz, Merced, San Benito and Monterey counties were clustered together representing the Central California Coast and Coast Ranges region. Sequences generated from individuals collected inside Fresno, Kern, Madera, and Mariposa counties were placed under Great Valley and Sierra Nevada region, whereas sequences from Los Angeles, San Bernardino and

San Diego were clustered into Southern California Coast, Mountain and Valley region. A Mantel test was performed with 999 randomized permutations to compare the genetic distances among *C. filaginifolia* individuals with corresponding geographic distances in GENALEX.

RESULTS

Amplification success varied across loci whereby ITS, *rpoC1*, and *ycf5* loci amplified in 95% of the samples; ITS2 amplified in 80% of the samples; *psbA-trnH* and *rbcL* amplified in 60% of the samples (Table 2). Amplicon lengths for each locus were comparable to sequences generated in previous studies using the same primer pairs (Kress et al. 2005; Chen et al. 2010) (Table 2).

Sequence Alignment, Partition, and Divergence

The overall alignment consistency score was 99 for each locus. The combined alignment of nrDNA loci (3'ETS, 5'ETS, and ITS) based on 31 sequences of *C. filaginifolia* consisted of 1884 characters. The best scheme showed only one partition for nrDNA dataset and K80+G DNA evolution model. The cpDNA alignment based on 20 sequences consisted of 2338 characters and the best scheme showed two partitions for this alignment. The first partition consisted of all cpDNA coding loci (*rbcL*, *rpoC1*, and *ycf5*) with F81+G DNA substitution model. The second partition consisted of the non-coding cpDNA gene (*psbA-trnH*), which was placed under HYG+G DNA evolution model. The nrDNA loci containing sequences from across California (3'ETS, 5'ETS and ITS) exhibited higher average divergences (K2P = 0.001–0.01) within *C. filaginifolia* in comparison to nrDNA loci (ITS2) and all cpDNA loci (K2P = 0.0002–0.001) which contained sequences generated from *C. filaginifolia* individuals sampled within San Diego County (Table 2).

Phylogenetic Analyses of nrDNA Loci

Maximum parsimony analysis showed 9 of 1884 characters to be parsimony informative for the ingroup samples. The average base frequencies were A = 0.22055; C = 0.23907; G = 0.26581; T = 0.27457. With the outgroup taxa added, the program retained 939 most parsimonious trees, and 941 tree islands were obtained from 1000 replications. The tree length of most parsimonious trees was 60. The rescaled consistency index and retention index were 0.9593 and 0.9756, respectively, which provide support for divergent evolution of sequences. The consistency index excluding uninformative characters was 0.9583 (Fig. 3a). Maximum likelihood analysis of the partitioned dataset yielded 1000 trees with the maximum likelihood score range of –3057.920 to –3057.928. The best tree was used to compute bipartitions with 1000 bootstrapped replicates, and

TABLE 1. GenBank accession numbers of *Corethrogyne flaginifolia* and outgroup sequences used in this study, generated from chloroplast DNA (*psbA-trnH*, *rbcL*, *rpoC1*, *ycf5*) and nuclear ribosomal DNA (ITS, and ITS2) loci. *Corethrogyne flaginifolia* var. *linifolia* (*C. f. var. linifolia*) leaf samples were collected from Torrey Pines State Natural Reserve (TPSR) extension in Del Mar, San Diego County, CA, and Point Loma (PL), San Diego County, CA. *Corethrogyne flaginifolia* var. *incana* (*C. f. var. incana*) leaf samples were collected from six different locations on Point Loma peninsula, San Diego County, CA, which included Cabrillo Drive (CabN, CabS), McClelland Road (McC), SPAWAR (Spa), Electron Drive (Elec), and National Park Service (NPS). The sample codes used in this study (i.e., Sample_ID column) include the location or county from where the sample was collected. The blank cells represent unsuccessful locus amplification in samples. All DNA samples are stored in J. Sharma's lab, Texas Tech University, Lubbock, TX.

Sample_ID	Taxon	Collector and Collection number	Herbarium accession number	Collection date	<i>psbA-trnH</i>	<i>rbcL</i>	<i>rpoC1</i>	<i>ycf5</i>	ITS	ITS2	3' ETS	5' ETS
CF_TP5R1	<i>C. f. var. linifolia</i>	19 May 2014	MF167349	MF167320	MF135577	MF167338	MF183976	MF183995	MF183975	MF167337	MF183975	MF183994
CF_TP5R2	<i>C. f. var. linifolia</i>	19 May 2014	MF167350	MF167321	MF135578	MF167339	MF183977	MF183996	MF183978	MF167336	MF183978	MF183997
CF_PL	<i>C. f. var. linifolia</i>	16 May 2014	MF167340	MF167310	MF135575	MF167336	MF183978	MF183997	MF183979	MF167322	MF183979	MF183998
CF_CabN1	<i>C. f. var. incana</i>	20 May 2014	MF167341	MF167311	MF135561	MF167323	MF183980	MF183999	MF183981	MF167324	MF183981	MF184000
CF_CabS2	<i>C. f. var. incana</i>	20 May 2014	MF167342	MF167312	MF135562	MF167324	MF183981	MF184000	MF183982	MF167325	MF183982	MF184001
CF_CabS3	<i>C. f. var. incana</i>	20 May 2014	MF167343	MF167313	MF135563	MF167325	MF183982	MF184001	MF183983	MF167329	MF183983	MF183984
CF_McC1	<i>C. f. var. incana</i>	20 May 2014			MF135567	MF167329	MF183984	MF183985	MF184002	MF167330	MF183985	MF184003
CF_Elec1	<i>C. f. var. incana</i>	20 May 2014			MF135568	MF167331	MF183986	MF184004	MF184005	MF167326	MF183986	MF184006
CF_Elec2	<i>C. f. var. incana</i>	20 May 2014			MF135565	MF167327	MF183987	MF184007	MF184008	MF167328	MF183988	MF184009
CF_Elec3	<i>C. f. var. incana</i>	20 May 2014	MF167344	MF167315	MF135566	MF167328	MF183989	MF184008	MF184009	MF167329	MF183989	MF184010
CF_Spa1	<i>C. f. var. incana</i>	20 May 2014			MF135572	MF167334	MF183990	MF184009	MF184010	MF167335	MF183991	MF184011
CF_Spa2	<i>C. f. var. incana</i>	20 May 2014			MF135573	MF167334	MF183990	MF184011	MF184012	MF167335	MF183991	MF184013
CF_Spa3	<i>C. f. var. incana</i>	20 May 2014	MF167348	MF167319	MF135574	MF167335	MF183990	MF184012	MF184013	MF167336	MF183991	MF184014
CF_NPS1	<i>C. f. var. incana</i>	20 May 2014	MF167345	MF167316	MF135569	MF167331	MF183991	MF184013	MF184014	MF167332	MF183992	MF184015
CF_NPS2	<i>C. f. var. incana</i>	20 May 2014	MF167346	MF167317	MF135570	MF167333	MF183992	MF184014	MF184015	MF167334	MF183993	MF184016
CF_NPS3	<i>C. f. var. incana</i>	20 May 2014	MF167347	MF167318	MF135571	MF167333	MF183993	MF184015	MF184016	MF167335	MF183994	MF184017
<i>Bidens pilosa</i>	<i>Bidens pilosa</i>	SD 152639	Reiser s.n.		GQ435068	GQ436430	GQ435995	GQ435601	GQ434471			
CF_San_Diego_1	<i>C. filaginifolia</i>	SD 152639	Reiser s.n.		SD 152639				JX680472.1			JX680499.1
CF_San_Diego_2	<i>C. filaginifolia</i>	SD 164291	Rebman 9514		SD 164291				JX680473.1			JX680500.1
CF_San_Diego_3	<i>C. filaginifolia</i>	SD 153413	Rich 109		SD 153413				JX680474.1			JX680501.1
CF_San_Diego_4	<i>C. filaginifolia</i>	SD 160694	Rebman 8252		SD 160694				JX680475.1			JX680502.1
CF_San_Diego_5	<i>C. filaginifolia</i>	SD 156442	Rebman 9579		SD 156442				JX680476.1			JX680503.1
CF_San_Diego_6	<i>C. filaginifolia</i>	SD 160377	Gregory 1179		SD 160377				JX680477.1			JX680504.1
CF_San_Diego_7	<i>C. filaginifolia</i>	JEPS 126512	Wallace s.n.		JEPS 126512				JX680478.1			JX680505.1
CF_Fresno_1	<i>C. filaginifolia</i>	JEPS 090954	Taylor 717		JEPS 090954				JX680479.1			JX680506.1
CF_Kern_1	<i>C. filaginifolia</i>	JEPS 126513	Markos 158		JEPS 126513				JX680480.1			JX680507.1
CF_San_Bernardino_1	<i>C. filaginifolia</i>	JEPS 126515	Markos 202		JEPS 126515				JX680481.1			JX680508.1
CF_San_Bernardino_2	<i>C. filaginifolia</i>	JEPS 126516	Markos 203		JEPS 126516				JX680482.1			JX680509.1
CF_Monterey_1	<i>C. filaginifolia</i>	JEPS 119402	CR 338		JEPS 119402				JX680483.1			JX680510.0

GenBank Accession numbers

TABLE 1. CONTINUED

Sample_ID	Taxon	Collector and Collection Number	Herbarium accession number	Collection date	GenBank Accession numbers							
					<i>psbA-trnH</i>	<i>rbcL</i>	<i>rpoC1</i>	<i>ycf5</i>	ITS	ITS2	3' ETS	5' ETS
Cf_Del_Norte_1	<i>C. filaginifolia</i>	<i>Parks & Tracy</i> <i>11496</i>	JEPS 15871						JX680484.1		JX680511.1	
Cf_Mendocino_1	<i>C. filaginifolia</i>	<i>Bacigalupi</i> <i>8993</i>	JEPS 35847						JX680485.1		JX680512.1	
Cf_Marin_1	<i>C. filaginifolia</i>	<i>Robbins 3891</i>	JEPS 22067						JX680486.1		JX680513.1	
Cf_Los_Angles_1	<i>C. filaginifolia</i>	<i>Markos 193</i>	JEPS 126514						JX680487.1		JX680514.5	
Cf_San_Diego_8	<i>C. filaginifolia</i>	<i>Moran 28334</i>	SD 105277						JX680488.1		JX680515.1	
Cf_Mariposa_1	<i>C. filaginifolia</i>	<i>Rose 60125</i>	JEPS 26049						JX680489.1		JX680516.1	
Cf_Santa_Clara_2	<i>C. filaginifolia</i>	<i>Ewan 8119</i>	UC 521768						JX680493.1		JX680520.1	
Cf_Merced_1	<i>C. filaginifolia</i>	<i>Lyon 1572</i>	UC 971972						JX680492.1		JX680519.1	
Cf_San_Benito_1	<i>C. filaginifolia</i>	<i>Lyon 1428</i>	UC 971971						JX680495.1		JX680522.1	
Cf_San_Bernardino_3	<i>C. filaginifolia</i>	<i>Jones 7239</i>	UC 1544816						JX680496.1		JX680523.0	
Cf_Monterey_2	<i>C. filaginifolia</i>	<i>Bacigalupi</i> <i>2692</i>	UC 708668						JX680497.1		JX680524.1	
Cf_San_Diego_10	<i>C. filaginifolia</i>	<i>Reveal 2754</i>	SD 110423						JX680498.1		JX680525.1	
Cf_Mendocino_3	<i>C. filaginifolia</i>	<i>Moore 1018</i>	JEPS 126518						JQ011998.1			
Cf_Santa_Cruz_1	<i>C. filaginifolia</i>	<i>Markos 116</i>	JEPS 88053						AF251593		AF251651	AF322271
Cf_Mendocino_2	<i>C. filaginifolia</i>	<i>Semple 8536</i>	UC 1557828						AF251594		AF251652	AF322272
Cf_Santa_Clara_1	<i>C. filaginifolia</i>	<i>Markos 112</i>	JEPS <i>s.n.</i>						AF251589		AF251647	AF322267
Cf_San_Diego_9	<i>C. filaginifolia</i>	<i>Markos 144</i>	JEPS 88048						AF251590		AF251648	AF322268
Cf_Kern_2	<i>C. filaginifolia</i>	<i>Markos 159</i>	JEPS 126511						AF251591		AF251649	
Cf_Madera_1	<i>C. filaginifolia</i>	<i>Markos 163</i>	JEPS 126510						AF251592		AF251650	AF322269
Lessingia_arachnoidea_1	<i>Lessingia arachnoidea</i>	<i>Markos 210</i>	JEPS 126509						JX680490.1		JX680517.1	
Lessingia_arachnoidea_2	<i>Lessingia arachnoidea</i>	<i>Markos 221</i>	JEPS 126517						JX680491.1		JX680518.1	
Lessingia_glandulifera_1	<i>Lessingia glandulifera</i>	<i>Twisselmann</i> <i>13730</i>	JEPS 54444						JX680494.1		JX680521.1	

TABLE 2. Amplification success of nuclear and chloroplast DNA barcoding loci used for resolving the phylogeny of *Corethrogyne filaginifolia*. Amplicon length represents the range of sequence length across all individuals of *C. filaginifolia*. Nuclear ribosomal DNA (nrDNA) loci 3' ETS, and 5' ETS represent GenBank extracted sequences of *C. filaginifolia*, while the ITS locus represents both GenBank (31), and self-generated (19) sequences. The ITS2 locus of nrDNA and all four chloroplast DNA (cpDNA) loci represent self-generated sequences. Pairwise sequence distances were calculated using Kimura 2-parameter (K2P) DNA substitution model with gamma distribution after pairwise deletion of gaps.

Locus	PCR success (%)	Sequencing success (%)	Amplicon length (bp)	n	Mean K2P distance	Range of K2P distances
nrDNA						
3' ETS	n/a	n/a	n/a	30	0.0031	0.000–0.012
5' ETS	n/a	n/a	n/a	5	0.0154	0.004–0.033
ITS	95	100	659–735	50	0.001	0.000–0.005
ITS2	80	100	436–463	16	0.0003	0.000–0.002
cpDNA						
<i>rbcL</i>	60	100	659–712	12	0.0002	0.000–0.001
<i>rpoC1</i>	95	100	506–510	19	0.0009	0.000–0.004
<i>ycf5</i>	95	100	383–402	19	0.0006	0.000–0.005
<i>psbA-trnH</i>	60	100	390–415	11	0.001	0.000–0.007

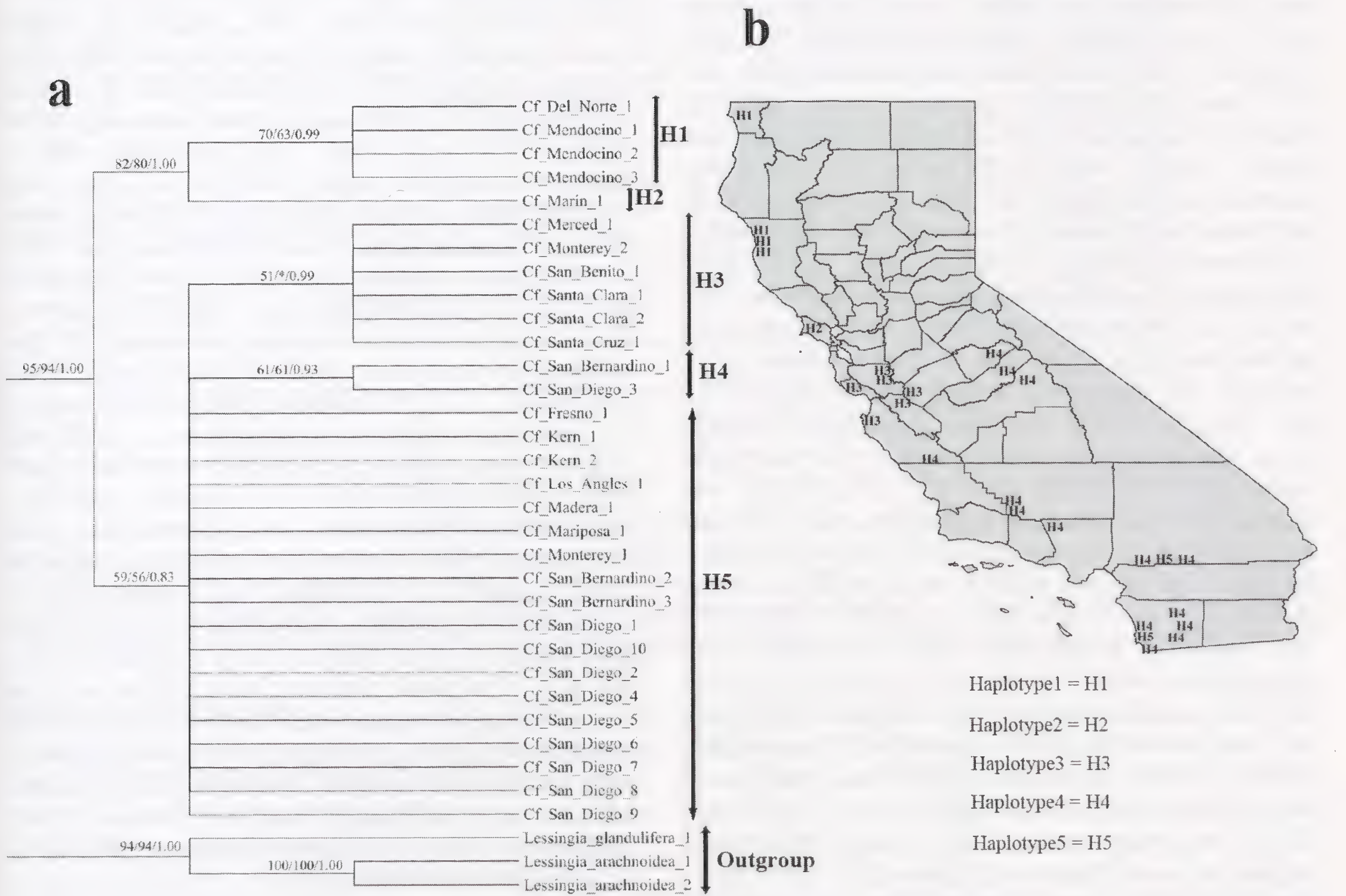


FIG. 3. a) Phylogeny of *Corethrogyne filaginifolia*. The tree was constructed using partitioned dataset that included previously generated GenBank reference sequences from ITS, 3'ETS, and 5'ETS loci from nuclear ribosomal DNA (nrDNA). The clade support values from maximum parsimony, maximum likelihood, and Bayesian analyses are shown for each clade (an * indicates <50% bootstrap support). The potential haplotype clades are indicated with text. b) California map showing the counties and locations of identified haplotypes, each sampling location is not visibly identifiable by text on the map because of the overlap.

TABLE 3. A summary of transitions in ITS and transversions in 3' ETS locus alignments among *Corethrogyne filaginifolia* individuals. The length of the aligned ITS locus was 630 characters and the length of 3' ETS locus was 565 characters. Transitions and transversions are in bold-faced font according to the characterization of different haplotypes. Haplotype5 is used as a reference character state due to its occurrence in majority of the sequences. The character (asterisk*) in haplotype2 differentiates it from haplotype1.

Individuals	Base position in alignment					
	ITS		3' ETS			
	388	619	240	273	325	396
Haplotype1	T	A	C	C	G	G
Haplotype2	T	G*	C	C	G	G
Haplotype3	C	G	G	A	G	T
Haplotype4	C	G	G	A	C	G
Haplotype5	C	G	G	A	G	G

the resultant consensus ML tree was obtained with <50–100% clade support values (Fig. 3a). Additionally, the Bayesian analysis showed convergence of two independent runs after 10 million generations with <0.01 average standard deviation of split frequencies, >200 Expected Sample Size (ESS) and >1 Potential Scale Reduction Factors (PSRF). Marginal likelihood scores of the two runs were very similar (–3081 and –3105) suggesting the search reached convergence. Of the 64 partitions, 30 partitions were informative with >0.1 probability of occurrence across the two runs (Fig. 3a). All three phylogenetic analysis methods were consistent with respect to the relative placement of *C. filaginifolia* individuals (Fig. 3a). The *C. filaginifolia* clade revealed five haplotypes based on ITS and 3'ETS loci (Fig. 3a, Table 3). These haplotypes resulted from two synapomorphies within ITS and four synapomorphies within the 3'ETS. Haplotype5 was chosen as the character reference state for the comparisons of synapomorphies (Table 3). The length of the aligned ITS locus was 630 characters and showed two transitions at the 388th and 619th base position. Haplotype1 and haplotype2 were distinguished based on these two transitions within the ITS locus. The length of the aligned 3'ETS locus was 564 characters, and there were four transversions at 240th, 273th, 325th, and 396th base positions in the alignment. The two transversions at 240th and 273th base positions in 3'ETS locus differentiated haplotype1 and haplotype2 from the other three haplotypes. The single transversions at 325th and 396th base position in 3'ETS locus distinguished haplotype3 and haplotype4, respectively. Additionally, the ITS tree including GenBank and sequences generated from this study grouped all *C. filaginifolia* individuals together to the exclusion of the outgroup samples. Nested within a largely unresolved clade of *C. filaginifolia* samples is a clade including individuals from haplotype1 and haplotype2 which is supported by moderate to high credibility values in

all three phylogenetic analyses (MP = 100%, ML = 63%, BI = 0.88) (Fig. 4).

Analysis of Molecular Variance (AMOVA) showed significant differences in nrDNA haplotype diversity of *C. filaginifolia* from four ecoregions ($p = 0.001$). The variation in haplotype diversity within ecoregions was 33%, whereas variation among ecoregions was 67%. The Mantel correlation between genetic and geographic distances was 0.21 ($p = 0.01$).

Phylogenetic Analyses of cpDNA Loci

Maximum Parsimony analysis in PAUP* showed seven out of 2338 DNA characters to be parsimony informative for cpDNA loci of the ingroup samples. The average base frequencies were A = 0.29771; C = 0.18650; G = 0.19061; T = 0.32518. After addition of outgroup taxon, numerous equally parsimonious trees were retained (911), and 915 tree islands were found from 1000 replications. The tree length of the most parsimonious trees was 307. The rescaled consistency index, and retention index were 0.9611 and 0.9643, respectively. The consistency index excluding uninformative characters was 0.9583 (Fig. 5a). One thousand ML trees generated in RAXML had maximum likelihood scores ranging from –4151.323 to –4152.060. The best tree, with a maximum likelihood score of –4151.323, was used to assign clade support values from 1000 bootstraps. Most of the clade support values were less than 50% for the ML analysis (Fig. 5c). In the Bayesian analysis, two independent runs had good convergence with the average standard deviation of split frequencies <0.01 after 2 million generations (ESS > 200, PSRF > 1). The marginal likelihood scores of the two runs were also homogenous (–4255 and –4256). A total 56 partitions were found from which 28 partitions were found to be informative with >0.1 probabilities of occurrence across the two runs (Fig. 5b). None of the analyses showed genetic discrimination between *C. filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia*. Unlike for the nrDNA data phylogenetic analyses, the topology was different for the each of the three methods. Also, the majority of the clade support values were <50% for ML and MP methods. ML, MP, and Bayesian analyses produced non-monophyletic varieties. In the cpDNA phylogenetic analyses, individual relationships do not reflect geography, nor do the relationships reflect any morphological similarities.

DISCUSSION AND CONCLUSIONS

We report five haplotypes of *C. filaginifolia* based on nrDNA barcoding data generated from specimens collected across California. Of the five haplotypes detected, haplotype1 and haplotype2 together segregated as a separate lineage from the rest (Fig. 3a). Among the five individuals constituting haplotype1 and haplotype2, three were from Mendocino County

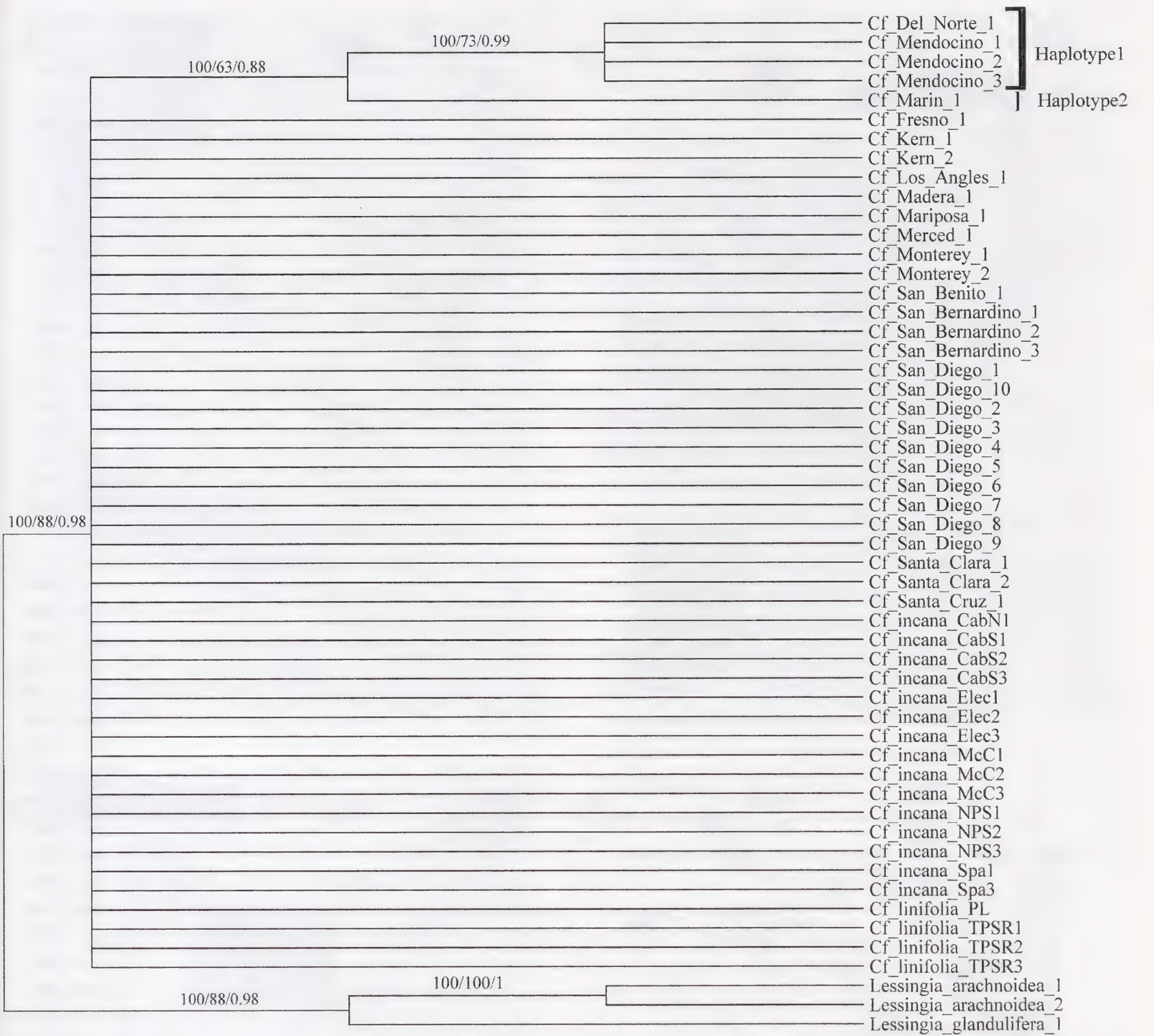


FIG. 4. Phylogeny of *Corethrogyne filaginifolia* based on ITS locus from nuclear ribosomal DNA (nrDNA). The tree was constructed using previously generated ITS GenBank reference sequences, and ITS sequence generated in this study. The clade support values from maximum parsimony, maximum likelihood, and Bayesian analyses are shown. The tree showed clustering of *C. filaginifolia* var. *incana*, *C. filaginifolia* var. *linifolia*, and other GenBank *C. filaginifolia* individuals except the clade of haplotype1 and haplotype2 individuals. *Lessingia arachnoidea* and *L. glandulifera* were used as outgroup species.

(*Bacigalupi* 8993 [JEPS], *Moore* 1018 [JEPS], and *Semple* 8536 [UC]) one from coastal Del Norte County (*Parks & Tracy* 11496 [JEPS]), and one from Marin County (*Robbins* 3891 [JEPS]) (Fig. 3a and b). The populations of *C. filaginifolia* in Mendocino, Del Norte and Marin counties experience similar environmental conditions given their placement within the Northern Coast section of California's coastal steppe, mixed forest and redwood forest province (McNab et al. 2007). It is possible that these populations constitute a different phylogeographic taxon of *C. filaginifolia*. Moreover, the support values for this clade in the nrDNA cladogram were high (MP = 82%, ML = 80%, Bayesian = 1.00) (Fig. 3a). Within this clade, an individual collected from Marin County (haplotype2) segregated from individ-

uals representing Del Norte and Mendocino counties (haplotype1) due to lack of shared transition at 619th base position in ITS locus (Table 3). However, we are reporting the five individuals belonging to haplotype1 and haplotype2 as a single phylogeographic taxon given their consistent transition and transversions at other positions in ITS and 3'ETS loci (Table 3). Further, the morphology of a specimen collected from Marin County (*Robbins* 3891 [JEPS]) was similar to the vouchers representing Del Norte (*Parks & Tracy* 11496 [JEPS]), and Mendocino counties (*Bacigalupi* 8993 [JEPS], *Moore* 1018 [JEPS], and *Semple* 8536 [UC]). Additional multiple samples of *C. filaginifolia* should be included from the southern limits of North Coast of California to document the genetic and morphological variation



FIG. 5. Phylogeny of *Corethrogyne filaginifolia* based on cpDNA, which included sequences generated in the present study from *psbA-trnH*, *rbcL*, *rpoC1*, and *ycf5* loci with partitioned dataset. The name of each individual is followed by the collection site code shown in Fig. 1 and Table 1. The trees provide topology and clade support values from, a) maximum parsimony (an * indicates <50% bootstrap support), b) Bayesian analyses, and c) maximum likelihood analyses methods. *Bidens pilosa* was added as an outgroup species. The trees showed polytomy of *C. filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia*.

within this region. The voucher specimens of haplotype1 (*Parks & Tracy 11496* [JEPS], *Bacigalupi 8993* [JEPS], *Moore 1018* [JEPS], and *Semple 8536* [UC]), and haplotype2 (*Robbins 3891* [JEPS]) showed clear morphological distinctions in foliage when compared to other haplotypes (Fig. 6). The leaves of haplotype1 and haplotype2 are spatulate and apically serrate (Fig. 6a), whereas the leaves of haplotypes 3, 4, and 5 are obovate, oblanceolate to linear, and without serrations (Fig. 6b–d). Haplotype3 was identified from specimens from Santa Cruz

(*Markos 116* [JEPS]), Santa Clara (*Ewan 8119* [UC]), San Benito (*Lyon 1428* [UC]), Monterey (*Bacigalupi 2692* [UC]), and Merced (*Lyon 1572* [UC]) counties, which fall under California central coast and coastal ranges ecoregions. The specimens from this region shared only a single transversion, and the clade support values for haplotype3 were low in both MP and ML analyses (51% and <50%, respectively) (Fig. 3a). Thus, we cannot be confident in haplotype3 being a separate phylogeographic taxon. Although, the morphology of specimens in haplotype3 was

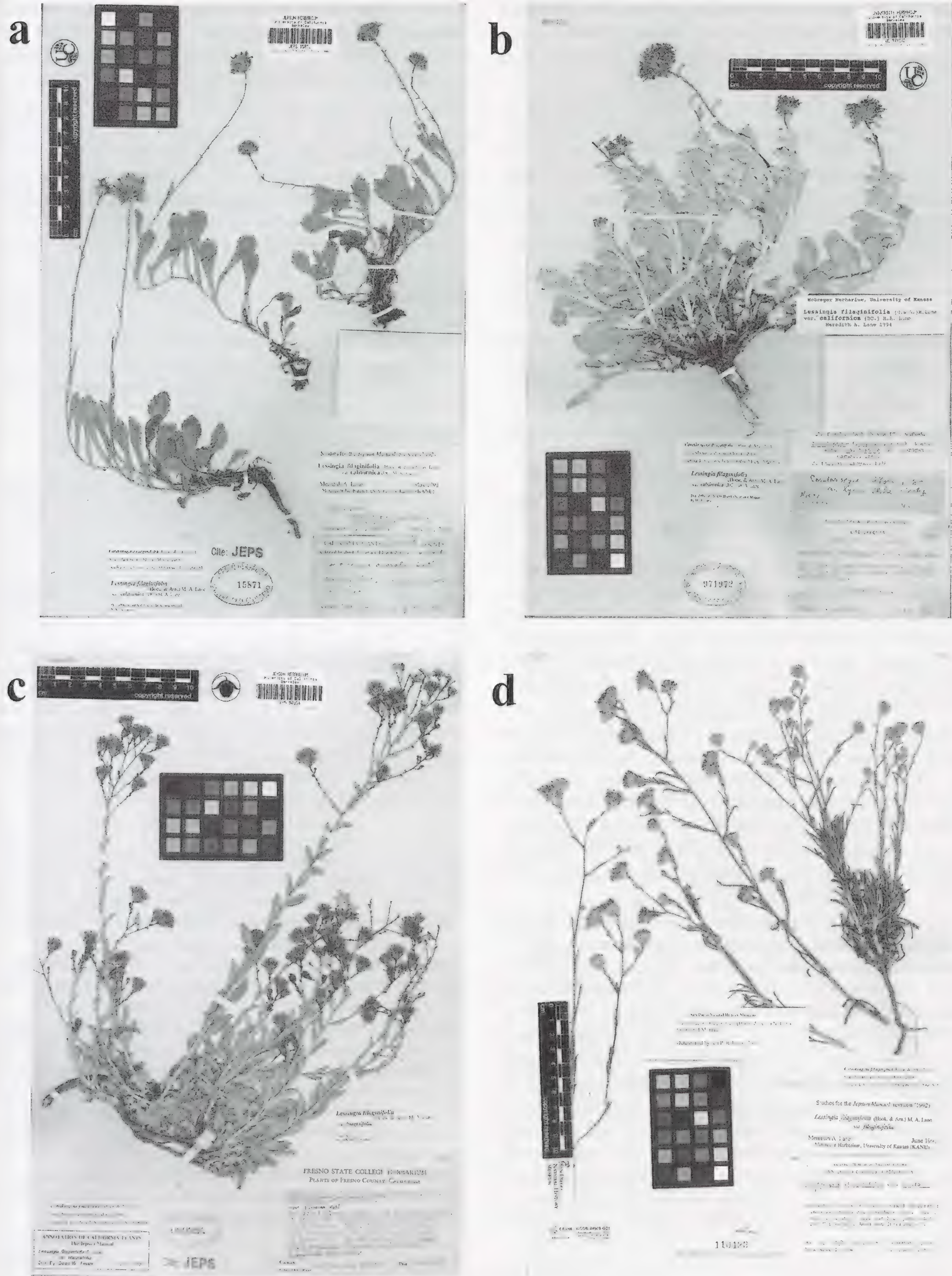


FIG. 6. Morphological variation among herbarium voucher specimens of *Corethrogyne filaginifolia* corresponding to the sequences from GenBank that we used as reference sequences to identify haplotypes. a) JEPS 15871, individual of *C. filaginifolia* from Del Norte County representing haplotype1 and haplotype2. The plant shows primary growth of shoot and spatulate leaves that are apically serrated. b) UC 971972, individual of *C. filaginifolia* from Merced County, a representative of haplotype3. The plant shows secondary growth with outward growth of shoots, and obovate leaves with entire margins. c) JEPS 090954, *C. filaginifolia* individual from Fresno County, a representative of haplotype5. The plants have oblanceolate leaves. d) SD 110423, *C. filaginifolia* var. *linifolia* individual from San Diego County, a representative of haplotype5 with narrow leaves and tomentose involucral bracts.

distinct in terms of plant growth, haplotype3 individuals exhibited more secondary growth along shoots in comparison to haplotype1, which showed mostly primary growth (Fig. 6a and b).

Haplotype4 was represented by two specimens, one from San Diego County (*Rich 109* [SD]), and the other from San Bernardino County (*Markos 202* [JEPS]). This haplotype was characterized by a single transversion in 3'ETS locus, and the clade support values were low in the parsimony and likelihood analyses (MP = 61%, ML = 61%), but high in the Bayesian analysis (BI = 0.93) (Fig. 3a). The representative specimen of haplotype4 from San Diego County (*Rich 109* [SD]) was identified as *C. filaginifolia* var. *linifolia*; whereas, the varietal annotation of the representative individual from San Bernardino County could not be ascertained (*Markos 202* [JEPS]). *Corethrogyne filaginifolia* var. *linifolia* has not been reported outside San Diego County, thus the grouping of *C. filaginifolia* var. *linifolia* individual from San Diego County with *C. filaginifolia* individual from San Bernardino County does not allow for confident inferences about haplotype4. Moreover, individual Cf_San_Diego_10 (*Reveal 2754* [SD]) from haplotype5 was also identified as *C. filaginifolia* var. *linifolia*. Haplotype5 was formed by individuals from Central Coast (*CR 338* [JEPS]), Great Valley and Sierra Nevada (*Taylor 717* [JEPS], *Markos 158* [JEPS], *Rose 60125* [JEPS], *Lyon 1572* [UC], *Markos 159* [JEPS], and *Markos 163* [JEPS]), Southern Coast, and Mountain Valley (*Reiser s.n.* [SD], *Rebman 9514* [SD], *Rebman 8252* [SD], *Rebman 9579* [SD], *Gregory 1179* [SD], *Wallace s.n.* [JEPS], *Markos 203* [JEPS], *Markos 193* [JEPS], *Moran 28334* [SD], *Jones 7239* [UC], and *Reveal 2754* [SD]) sections of California's ecoregions. The polytomy including haplotype5 could not be resolved despite its distinctive morphological variation, especially in San Diego County (Fig. 3a).

We did not detect genetic differences among individuals of *C. filaginifolia* var. *incana* and *C. filaginifolia* var. *linifolia* which were examined with both cpDNA and nrDNA data. The distribution of individuals of the two varieties within nrDNA (Fig. 4) and cpDNA (Fig. 5) cannot be explained with respect to either their provenance or morphological characteristics. This result either implies the occurrence of varying rates of mutation in the two varieties at these loci or the unsuitability of cpDNA and nrDNA loci to accurately explain the infraspecific variation for these two varieties. The discordance between gene trees and morphological distinctions of the two varieties can also be a result of incomplete lineage sorting.

The genetic distinction of northern *C. filaginifolia* individuals (haplotype1 and haplotype2) from the rest of the individuals at nrDNA loci supports the morphometric study of *C. filaginifolia* by Saroyan et al. (2000), and we provide evidence in our study to classify these individuals as a distinct variety. The Mantel correlation ($r^2 = 0.21$, $p = 0.01$) between

genetic and geographic distances also provided evidence that morphological variation might be linked to genetic variation along the latitudinal gradient in *C. filaginifolia*. Although morphometric differences also exist within and between other haplotypes of *C. filaginifolia*, the genetic analyses with universal plant barcoding loci could only provide strong evidence for haplotype1 and haplotype2. We suspect that it might not be possible to fully resolve the infraspecific variation in *C. filaginifolia* using the standard plant barcoding loci. Although higher resolution polymorphic markers such as double digest restriction site associated DNA sequencing (ddRAD-seq) of genomic simple sequence repeats (SSRs), single nucleotide polymorphism (SNPs), and transposable elements are some potential approaches to differentiate the infraspecific variability within *C. filaginifolia*.

ACKNOWLEDGMENTS

We gratefully acknowledge the Office of Naval Research for supporting this research through Cooperative Agreement No. N62473-13-2-4907 on behalf of the Naval Base Point Loma (U.S. Department of Defense). We also acknowledge National Park Service for allowing plant material collections under collection permit No. CABR-2015-SCI-0010. We are thankful to Dr. Staci Markos and, Dr. A. J. Moore for their contributions that included previously generated sequences and recent correspondence. Additionally, we thank Dr. John L. Strother for providing helpful comments on the early drafts of the manuscript, and Andrew S. Doran (University and Jepson Herbaria) and Layla Aerne (San Diego Natural History Museum) for providing high resolution images of voucher specimens.

LITERATURE CITED

- BALDWIN, B. G., D. H. GOLDMAN, D. J. KEIL, R. PATTERSON, T. J. ROSATTI, AND D. H. WILKEN (eds.). 2012. The Jepson manual: vascular plants of California, second edition. University of California Press, Berkeley, CA.
- BENSON, D. A., M. S. BOGUSKI, D. J. LIPMAN, J. OSTELL, B. F. OUELLETTE, B. A. RAPP, AND D. L. WHEELER. 1999. GenBank. Nucleic Acids Research 27:12–17.
- THE CALFLORA DATABASE (CALFLORA). 2015. Information on California plants for education, research and conservation, with data contributed by public and private institutions and individuals, including the Consortium of California Herbaria. Berkeley, CA. Website <http://www.calflora.org> (accessed 24 February 2015).
- CBOL PLANT WORKING GROUP. 2009. A DNA barcode for land plants. Proceedings of the National Academy of Sciences USA 106:12794–12797.
- CHASE, M. W., R. S. COWAN, P. M. HOLLINGSWORTH, C. VAN DEN BERG, S. MADRIÑÁN, G. PETERSEN, O. SEBERG, T. JØRGENSEN, K. M. CAMERON, AND M. CARINE. 2007. A proposal for a standardized protocol to barcode all land plants. Taxon 56:295–299.
- CHEN, S., H. YAO, J. HAN, C. LIU, J. SONG, L. SHI, Y. SHU, X. MA, T. GAO, AND X. PANG. 2010. Validation of the *its2* region as a novel dna barcode for identifying medicinal plant species. PLoS ONE 5:e8613.

- CALIFORNIA NATIVE PLANT SOCIETY (CNPS). 2015. Inventory of Rare and Endangered Plants (online edition, v8-02). California Native Plant Society, Sacramento, CA. Website <http://www.rareplants.cnps.org> (accessed 02 March 2015).
- DARRIBA, D., G. L. TABOADA, R. DOALLO, AND D. POSADA. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9:772.
- FERRIS, R. S. 1960. Bignoniaceae to Compositae: Bignonias to Sunflowers. Pp. 337–342 in L. R. Abrams and R. S. Ferris (eds.), Vol IV, An illustrated flora of the Pacific States: Washington, Oregon, and California (in four vols.). Stanford University Press, Stanford, CA.
- FUNK, V. A., A. SUSANNA, T. F. STUESSY, AND R. J. BAYER, Editors. 2009. Systematics, Evolution, and Biogeography of Compositae. International Association for Plant Taxonomy, Vienna.
- GAO, T., H. YAO, J. SONG, C. LIU, Y. ZHU, X. MA, X. PANG, H. XU, AND S. CHEN. 2010. Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2. *Journal of Ethnopharmacology* 130:116–121.
- HEBERT, P. D., S. RATNASINGHAM, AND J. R. DE WAARD. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological Sciences* 270(Suppl 1):S96–S99.
- JANSEN, R. K. AND J. D. PALMER. 1987. A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). *Proceedings of the National Academy of Sciences* 84: 5818–5822.
- KECK, D. D. 1959. Compositae. Pp. 1204–1207 in P. A. Munz and D. D. Keck (eds.), *A California flora*. University of California Press, Berkeley, CA.
- KRESS, W. J., K. J. WURDACK, E. A. ZIMMER, L. A. WEIGT, AND D. H. JANZEN. 2005. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* 102:8369–8374.
- KRESS, W. J. AND D. L. ERICKSON. 2007. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE* 2:e508.
- LANE, M. A. 1992. New combinations in Californian *Lessingia* (Compositae: Astereae). *Novon* 2:213–214.
- LANE, M. A. 1993. *Lessingia*. Pp 304–306 in J. C. Hickman (ed.), *The Jepson manual: higher plants of California*. University of California Press, Berkeley, CA.
- LANFEAR, R., B. CALCOTT, S. Y. HO, AND S. GUINDON. 2012. PartitionFinder version 1.1.1: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* 29:1695–1701.
- NCBI POPULATION SET (POPSET) 417346869. 2012. Intra-specific variation in *Corethrogyne filaginifolia* (Asteraceae). Website https://www.ncbi.nlm.nih.gov/popset/DbFrom=nucore&Cmd=Link&LinkName=nucore_popset&IdsFromResult=417346879 (accessed 10 November 2015).
- MARKOS, S. AND L. STROTHER. 2016. *Corethrogyne filaginifolia*, in Jepson Flora Project (eds.) *Jepson eFlora*. Website http://ucjeps.berkeley.edu/cgi-bin/get_IJM.pl?tid=2375, (accessed 04 February 2016).
- MARKOS, S. AND B. G. BALDWIN. 2001. Higher-level relationships and major lineages of *Lessingia* (compositae, astereae) based on nuclear rdna internal and external transcribed spacer (ITS and ETS) sequences. *Systematic Botany* 26:168–183.
- . 2002. Structure, molecular evolution, and phylogenetic utility of the 5' region of the external transcribed spacer of 18S-26S rDNA in *Lessingia* (Compositae, Astereae). *Molecular Phylogenetics and Evolution* 23:214–228.
- MCNAB, W. H., D. T. CLELAND, J. A. FREEOUF, G. J. NOWACKI, AND C. A. CARPENTER. 2007. Description of ecological subregions: sections of the conterminous United States. Gen. Tech. Report WO-76B. Washington, DC: U.S. Department of Agriculture, Forest Service. 80 p.
- MORGAN, D. R. 1997. Reticulate evolution in *Machaeranthera* (Asteraceae). *Systematic Botany* 22:599–615.
- MUNZ, P. A. 1974. *A flora of southern California*. University of California Press, Berkeley, CA, USA.
- NESOM, G. L. 1994. Subtribal classification of the Astereae (Asteraceae). *Phytologia* 76:193–274.
- NOTREDAME, C., D. G. HIGGINS, AND J. HERINGA. 2000. T-Coffee version 11.00: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* 302:205–217.
- PANERO, J. L. AND V. A. FUNK. 2008. The value of sampling anomalous taxa in phylogenetic studies: major clades of the Asteraceae revealed. *Molecular Phylogenetics and Evolution* 47:757–782.
- PEAKALL, R. O. D. AND P. E. SMOUSE. 2006. GENALEX version 6.5: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288–295.
- RAMBAUT, A. 2007. FigTree version 1.4.2: a graphical viewer of phylogenetic trees. <http://tree.bio.ed.ac.uk/software/figtree/> (accessed 10 January 2016)
- RAMBAUT, A. AND A. J. DRUMMOND. 2007. Tracer version 1.6 [computer program]. Website <http://tree.bio.ed.ac.uk/software/tracer/> (accessed 12 January 2016).
- RONQUIST, F. AND J. HUELSENBECK. 2003. MrBayes version 3.2.5: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
- SAROYAN, J. P., D. R. PARNELL, AND J. L. STROTHER. 2000. Revision of *Corethrogyne* (Compositae: Astereae). *Madroño* 47:89–96.
- STAMATAKIS, A. 2006. RAXML-VI-HPC version 7.4.2: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
- STUCKY, B. J. 2012. SeqTrace: a graphical tool for rapidly processing DNA sequencing chromatograms. *Journal of Biomolecular Techniques* 23:90–93.
- SWOFFORD, D. L. 2003. PAUP* version 40a146: Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates Inc., Sunderland, MA, USA.
- TAMURA, K., G. STECHER, D. PETERSON, A. FILIPSKI, AND S. KUMAR. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30:2725–2729.
- THOMPSON, J. D., F. PLEWNIK, AND O. POCH. 1999. BALiBASE: a benchmark alignment database for the evaluation of multiple alignment programs. *Bioinformatics* 15:87–88.
- ZHANG, X. AND K. BREMER. 1993. A cladistic analysis of the tribe Astereae (Asteraceae) with notes on their evolution and subtribal classification. *Plant Systematics and Evolution* 184:259–283.